知识管理论坛 ISSN 2095-5472 CN11-6036/C
Knowledge Management Forum

E-mail: kmf@mail.las.ac.cn
http://www.kmf.ac.cn

【学术探索】

在线健康社区中基于用户属性的时序交互模式研究

◎ 吴冰 彭彧

同济大学经济与管理学院 上海 200092

摘要：[目的 / 意义] 在线健康社区迅速发展，但缺乏结合网络结点属性与网络结构的动态网络特征的研究，因而难以揭示基于结点属性的用户动态交互模式的形成机理。[方法 / 过程] 应用基于节点属性的时序指数随机图模型，以在线健康社区为研究对象，结合用户节点属性特征，包括用户发文情感倾向、用户发文文本长度和用户社区等级，构建在线健康社区的用户时序交互模式研究模型，从百度贴吧的糖尿病吧抓取 2016 年 10 月至 2018 年 2 月期间 2 301 个有效用户，6 045 条主帖和 9 490 条回复，实证用户时序交互模式特征。[结果 / 结论] 用户属性特征对互惠性时序模式、k-star 时序模式、传递性时序模式和循环性时序模式形成有显著影响，并由此为在线健康社区建设提出发展建议。

关键词：在线健康社区 节点属性 用户交互模式 NATERGM

分类号：C912

引用格式：吴冰，彭彧. 在线健康社区中基于用户属性的时序交互模式研究 [J/OL]. 知识管理论坛，2019, 4(3): 163-172[引用日期]. http://www.kmf.ac.cn/p/174/.

在线健康社区以医疗信息分享和交流为主要目的，越来越多的人尝试通过在线健康社区解决健康相关问题。因此，在线健康社区的发展正在推动传统医疗改革，以逐步缓解我国医疗资源供需失衡 [1]。根据用户的构成，在线健康社区可以分为 3 类：医疗工作者社区（如丁香园）、医患社区（如好大夫在线）和病患者社区（如百度医疗贴吧），分别为医疗工作者、医患和病患者之间相互发现与相互支持搭建了沟通与联系的桥梁。近年来，我国慢性病患者持续增长，以糖尿病及其并发症为例，据国际糖尿病联盟（International Diabetes Federation, IDF）报告显示，预计到 2045 年，全球糖尿病患者可能达到 6.29 亿。慢性病群体更倾向于利用互联网来搜寻疾病信息以及与患者交流 [2]，因此，以病患者之间互助为中心的病患者社区成为病患者信息互惠、经验分享与情感支持的重要平台，尤其需要揭示用户属性与时序交互模式形成的内在机理，但目前相关研究缺乏。

因此，笔者将从动态集成视角，应用基于节点属性的时序指数随机图模型（Node Attribute-based Temporal Exponential Random

作者简介：吴冰（ORCID:0000-0002-5936-1631），副教授，博士，E-mail：ww_bing@163.com；彭彧（ORCID: 0000-0002-4335-7929），硕士研究生。

收稿日期：2019-02-27 发表日期：2019-06-18 本文责任编辑：刘远颖

Graph Model ,NATERGM)，结合在线病患者社区的用户节点属性特征，构建在线病患者社区的用户时序交互模式研究模型，从百度贴吧糖尿病吧抓取数据，进行实证研究，探究基于用户节点属性的动态交互模式形成机理。在理论上，丰富和扩展在线健康社区的用户时序交互机制研究体系；在实践上，为在线健康社区的建设与发展提供方向和建议。

1 相关文献综述

1.1 在线健康社区的用户交互模式分析

在线健康社区被定义为关注共同健康问题的用户集合，用户可以被视为行动者，用户之间的交互可以被视为不同行动者之间的关联。目前，关于在线健康社区的用户交互模式分析，主要有 3 类研究：①从社区网络结构角度，采用社会网络分析，描述特定的在线健康社区中社会网络结构特点及其影响因素。例如，社会网络分析得到的中心性可以作为挖掘在线健康社区中关键用户的主要特征之一 [3]。②从社区网络个体角度，研究关注在线健康社区中个体属性及其行为模式。例如，在线健康社区中知识分享网络具有小世界效应，社区成员最关注他人分享的经验知识，不同用户群体的知识分享行为和活跃时长有显著差异 [4]；社会网络分析用户行为，可以探索在线健康社区网络中不同群组间以及不同角色间的知识分享行为 [5-6]。③从社区网络形成机理角度，应用指数随机图模型(Exponential Random Graph Model, ERGM)，研究在线健康社区用户交互模式。例如，利用 ERGM 分析在线健康社区中用户属性特征和行为特征对网络连接模式的影响 [7-9]。

由此，目前在线健康社区中用户交互模式的相关研究主要存在两个问题：①孤立分析网络结点属性与网络结构，然而网络结点属性是用户交互网络结构的重要组成；②缺乏网络结点属性与用户交互网络的动态集成研究，因而难以揭示网络结点属性与动态网络连接形成的内在机理。

在时序指数随机图模型(Temporal Exponential Random Graph Model，TERGM)基础上，近年来发展的基于节点属性的时序指数随机图模型将结点属性整合到 TERGM，由此该模型适用于研究节点属性作用下的用户交互网络动态形成机理。例如，从知识共享社区收集实证数据，应用 NATERGM 可以识别和预测节点属性对网络二元和三元关系动态形成模式的影响 [10]。因此，笔者将应用 NATERGM 研究在线健康社区中节点属性对用户动态交互模式的影响，揭示用户时序交互模式形成机理。

1.2 NATERGM

给定一个观察网络，指数随机图模型的主要任务是统计检验网络中哪些节点连接模式的出现不是偶然，如公式(1)所示 [10]：

$$P(Y=y)=(\frac{1}{k})\exp\{\sum_{A}\eta_A g_A(y)\} \quad \text{公式(1)}$$

其中：① $P(Y=y)$ 是一个概率分布，Y 是一个随机变量矩阵，表示网络连接，y 是 Y 的实现；② A 是节点连接模式，$g_A(y)$ 是与 A 相对应的网络统计；③ k 是一个比例参数，确保 $P(Y=y)$ 是一个随机分布；④ η_A 是与节点连接模式 A 相对应的参数，与节点连接模式 A 出现的可能性正相关，如果参数 η_A 是显著的，表明节点连接模式 A 更有可能会出现在网络中，由此说明节点连接模式 A 对网络形成过程有重要的影响。

TERGM 是 ERGM 向动态网络的延伸，在公式(1)的基础上，一阶马尔可夫依赖的 TERGM，如公式(2)所示 [10]，TERGM 的主要目标是检测在时序(t–1 和 t)，节点连接时序模式 A 是否更可能出现。

$$P(Y^t=y^t|Y^{t-1}=y^{t-1})=(\frac{1}{k(y^{t-1})})\exp\{\sum_{A}\eta_A g_A(y^t,y^{t-1})\} \quad \text{公式(2)}$$

NATERGM 是在 TERGM 中增加节点属性，在公式 (2) 的基础上，NATERGM 的计算如公式(3)所示 [10]：

知识管理论坛，2019(3):163–172
DOI：10.13266/j.issn.2095–5472.2019.017

$$P(Y=y|\eta)=(\frac{1}{k})\exp\left\{\sum_{a\in A}\eta_a g_a(y,T,X)\right\}$$

公式（3）

其中：① A 是要验证的节点连接时序模式集合；② $\eta=[\eta_a]$ 是一个向量参数，表示每个节点连接时序模式对网络形成的影响；③ $g_a(y, T, X)$ 是节点连接时序模式 a 的网络统计，参数为：网络 y、时间戳矩阵 T 和节点属性向量 X；④参数估计采用马尔可夫蒙特卡罗（Markov Chain Monte Carlo，MCMC）方法。因此，NATERGM 的主要目标是检验具有属性 X 的节点在时序 T，节点连接时序模式 A 是否更可能出现。

2 在线健康社区中用户时序交互模式的 NATERGM 分析框架

NATERGM 通过检查带有时间戳的网络节点连接截面数据，准确地跟踪网络节点连接时序，其基本分析框架如图 1 所示：

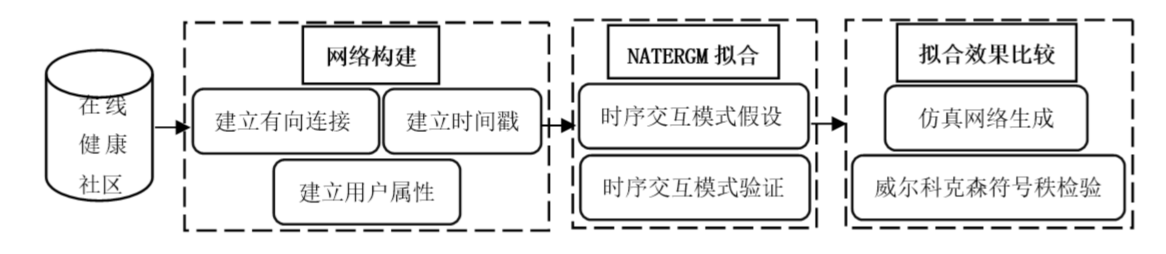


图 1 NATERGM 分析框架

2.1 网络构建

首先，基于在线健康社区用户的交互关系构建有向网络连接，用 Y_{ij} 表示网络用户节点 i,j 之间的有向连接，如果节点 i 指向节点 j，则 $Y_{ij}=1$；否则 $Y_{ij}=0$。识别用户节点间所有的有向连接后，具有 N 个节点的有向网络连接可表示为矩阵 $Y=[Y_{ij}], (i,j=1,2,\cdots, N)$。

其次，基于在线健康社区用户交互的有向网络连接，用 T_{ij} 记录网络用户节点 i,j 之间的有向连接建立时间，如果节点 i 指向节点 j 的时间早于节点 j 指向节点 i 的时间，则 $T_{ij}<T_{ji}$。记录用户节点间所有的有向连接建立时间后，具有 N 个节点的有向网络连接时间可以表示为矩阵 $T=[T_{ij}], (i,j=1,2,\cdots, N)$。

第三，基于在线健康社区用户的节点属性特征，用 x_i 表示每个网络节点 i 的属性值，具有 N 个节点的有向网络节点属性值，可表示为向量 $X=[x_1,x_2,\cdots,x_N], (i,j=1,2,\cdots,N)$。根据在线健康社区的社会支持作用，笔者将用户发文情感倾向、用户发文文本长度和用户社区等级作为用户属性向量，原因在于：①在线健康社区中，情感因素对病患交互的形成起着至关重要的作用 [11]；②平均发帖长度更长，具有较高写作能力的用户有助于提供更多的信息支持 [12]；③在线健康社区中，不同的用户行为，如每日签到、发帖和回复等，都被赋予了大小不等的经验值，通过累计经验等级，体现了用户的参与度和贡献度。

2.2 时序交互模式的 NATERGM 拟合

根据在线健康社区的用户属性向量，笔者着重研究二元关系 [13] 的互惠性以及三元关系 [13] 的 k-star 关系、传递性和循环性，见表 1- 表 4。其中，白节点代表一般用户，黑节点代表具有某一属性特征的用户；虚线箭头连接是在实线箭头连接基础上发展的网络连接；当节点属性 X_i 为分类属性，$I(X_i)$=0 或 1，当节点属性 X_i 为连续变量，$I(X_i)$ 为 X_i 属性值。

2.2.1 互惠性时序模式

互惠性（reciprocity）是指在网络中两个互相连接的节点所构成的子集，可以描述为

$\{(j\rightarrow i), (i\rightarrow j)\}$。在线社区的本质是一个基于社会资本交换和文化互惠性的实体 [14]，在线健康社区中用户交互具有互惠性 [7-8]。由此，笔者将结合时序，研究节点属性对互惠性时序连接产生的影响，时序假设及其网络统计公式，如表 1 所示：

表 1 NATERGM 的时序互惠性

互惠模式	图示	时序假设
反馈 (feedback)		H1a：发文情感积极的用户更可能收到其他用户的反馈 H1b：发文文本较长的用户更可能收到其他用户的反馈 H1c：社区等级较高的用户更可能收到其他用户的反馈
回复 (Response)		H2a：发文情感积极的用户更可能回复其他用户 H2b：发文文本较长的用户更可能回复其他用户 H2c：社区等级较高的用户更可能回复其他用户

2.2.2 k-star 时序模式

k-star 模式指网络中某个节点和其他数个与其相连的节点所组成的子集，笔者重点研究 2-star 模式。在有向图中，k-star 被分为 k-in-start 和 k-out-star。k-in-star 在一定程度上反应了节点在网络中受欢迎的程度或者聚敛性；k-out-star 在一定程度上反应了节点在网络中的发散性或者社会影响力，网络节点的受欢迎程度和社会影响力会影响网络连接的形成 [15]，在线健康社区中用户交互存在 k-star 连接 [7]。由此，笔者将结合时序，研究节点属性对 k-star 时序连接产生的影响，时序假设及其网络统计公式，如表 2 所示：

表 2 NATERGM 的时序 k–star

交互模式	图示	时序假设
k-in-star 抢先 (Initial)		H3a：发文情感积极的用户更可能抢先连接其他用户 H3b：发文文本较长的用户更可能抢先连接其他用户 H3c：社区等级较高的用户更可能抢先连接其他用户
k-in-star 拖延 (Laziness)		H4a：发文情感积极的用户更可能拖延连接其他用户 H4b：发文文本较长的用户更可能拖延连接其他用户 H4c：社区等级较高的用户更可能拖延连接其他用户
k-out-star 优先 (Priority)		H5a：发文情感积极的用户更可能被其他用户优先连接 H5b：发文文本较长的用户更可能被其他用户优先连接 H5c：社区等级较高的用户更可能被其他用户优先连接
k-out-star 滞后 (DePriority)		H6a：发文情感积极的用户更可能被其他用户滞后连接 H6b：发文文本较长的用户更可能被其他用户滞后连接 H6c：社区等级较高的用户更可能被其他用户滞后连接

2.2.3 传递性时序模式

对于网络中 i,j,k 三个节点，传递性（transitivity）可以被描述为边集合：$\{(i\rightarrow j), (i\rightarrow k), (j\rightarrow k)\}$，笔者研究传递性桥接、传递性追随和传递性引用。在社会网络中，人们倾向于和朋友的朋友建立网络关系 [16]，在线健康社区中用户交互存在传递性连接关系 [7]。由此，笔者将结合时序，研究节点属性对传递性时序连接所产生的影响，时序假设及其网络统计公式如表 3 所示：

知识管理论坛，2019(3):163–172
DOI：10.13266/j.issn.2095–5472.2019.017

表 3 NATERGM 的传递性时序

交互模式	图示	时序假设
桥接 (bridge)		H7a：发文情感积极的用户更可能桥接其他用户 H7b：发文文本较长的用户更可能桥接其他用户 H7c：社区等级较高的用户更可能桥接其他用户
追随（followup）		H8a：发文情感积极的用户更可能传递追随其他用户 H8b：发文文本较长的用户更可能传递追随其他用户 H8c：社区等级较高的用户更可能传递追随其他用户
引用（reference）		H9a：发文情感积极的用户更可能被其他用户传递引用 H9b：发文文本较长的用户更可能被其他用户传递引用 H9c：社区等级较高的用户更可能被其他用户传递引用

2.2.4 循环性时序模式

对于网络中 i,j,k 三个节点，循环性（cyclicity）可以被描述为边集合：$\{(i\rightarrow j), (j\rightarrow k), (k\rightarrow i)\}$，与传递性研究对应，笔者研究循环性反向桥接、循环性反向追随和循环性反向引用。在对狒狒族群所构成的社会网络的研究过程中发现，当食物短缺时，个体间表现出了明显的循环性交互 [17]，在线健康社区中用户交互存在循环连接关系 [7]。因此，笔者将结合时序，研究节点属性对循环性时序连接产生的影响，时序假设及其网络统计公式如表 4 所示：

表 4 NATERGM 的时序循环性

交互模式	图示	时序假设
反向桥接 (reversed bridge)		H10a：发文情感积极的用户更可能反向桥接其他用户 H10b：发文文本较长的用户更可能反向桥接其他用户 H10c：社区等级较高的用户更可能反向桥接其他用户
反向追随（reversed followup）		H11a：发文情感积极的用户更可能循环追随其他用户 H11b：发文文本较长的用户更可能循环追随其他用户 H11c：社区等级较高的用户更可能循环追随其他用户
反向引用（reversed reference）		H12a：发文情感积极的用户更可能被其他用户循环引用 H12b：发文文本较长的用户更可能被其他用户循环引用 H12c：社区等级较高的用户更可能被其他用户循环引用

2.3 NATERGM 拟合效果比较

比较步骤如下：首先，基于在时间点 t–1 观测的实际网络，估计 NATERGM 的参数 η_{t-1}；接着，基于 NATERGM 仿真生成时间点 t 的 K 个与实际类似的网络；第三，将仿真生成的 K 个网络与实际网络比较绝对差异，如公式(4)所示：

$$AD_a^t = \left| g_a(y_0^t) - \left(\frac{1}{K}\sum_{k=1}^{K} g_a(y_k^t)\right) \right| \qquad 公式(4)$$

其中，y_0^t 是时间点 t 的观测的实际网络，$y_k^t(k=1,2,\ldots,K)$ 是基于时间点 t 拟合模型的第 k 个生成网络，$g_a(y_k^t)$ 是节点连接时序模式 a 的网络；第四，由于 NATERGM 是在 TERGM 时序模式的基础上加入了节点属性，因而笔者将 TERGM 作为基线模型 (baseline model)，用于比较 NATERGM 拟合效果，采用威尔科克森符号秩检验验证 NATERGM 的拟合误差是否显著性低于 TERGM 的拟合误差，如公式（5）所示 [10]：

$$Z=(\left|W^{+}-\mu_{W}\right|-0.5)/\sigma_{W} \qquad 公式（5）$$

其中：① $W^{+}=\sum_{t=1}^{T}I(AD_{baseline}^{t}-AD_{NATERGM}^{t}>0)\cdot R_{i}$，$T$ 是比较周期，$I()$ 是基线模型 TERGM 拟合误差大于 NATERGM 的指示函数，R_i 是第 i 个比较对的秩；② $\mu_W=n(n+1)/4$；③ $\sigma_{W}=\sqrt{n(n+1)(2n+1)/24}$；④ Z 值越高说明 NATERGM 拟合误差显著性低于 TERGM 拟合误差。

3 百度糖尿病贴吧实证研究

自 2003 年创建以来，尽管百度贴吧并非专业的在线健康社区，但其以众多优点吸引了大量的病患，组成了庞大的病患者贴吧群，目前注册用户超过 10 亿，月活跃用户突破 3 亿，共拥有 1 900 万个主题吧。笔者选取百度糖尿病贴吧作为最终的研究对象，主要原因在于：糖尿病是一种较常见的疾病，病患人群广，病患更倾向于寻求情感慰藉和信息支持。

3.1 用户交互网络构建

3.1.1 网络连接

由于百度贴吧会定期清理冗余数据，因此抓取数据的时间跨度为 2016 年 10 月 -2018 年 2 月，共计 66 周。笔者采用 Python 的 Scrapy 框架，在抓取的时间段内，百度糖尿病吧中共有 23 777 个用户、77 970 条主帖和 80 729 条回复，其中：① 主帖是由用户直接发出在社区中广播，按照时间顺序出现在社区主页，是建立用户连接的一种邀请；② 回复是在主帖下的讨论区，通过自由回复建立用户连接。

数据抓取后，通过以下步骤进行清洗和预处理：① 删除没有任何回复的主帖；②删除没有参与任何交互的用户；③删除只有视频 / 表情包 / 图片的帖子；④删除时间相近的重复发帖。预处理后得到 2 301 个有效用户、6 045 条主帖和 9 490 条回复。

3.1.2 节点属性

(1) 用户发文的情感倾向。基于中文情感极性词典 NTUSD，第 i 个用户第 j 条发文的情感评分 $senti(p_{ij})$，如公式（6）所示 [18]：

$$Senti(p_{ij})=\frac{1}{\left|p_{ij}\right|}\sum_{s\in p_{ij}}[\delta(s)\times\sum_{w\in s}ntusd(w,pol(w))] \qquad 公式（6）$$

其中：① $|p_{ij}|$ 是第 i 个用户第 j 条发文的句子数；② s 是第 i 个用户第 j 条发文 p_{ij} 中的语句；③如果语句 s 是情感非中立的，$\delta(s)=1$，如果语句 s 是情感中立的，$\delta(s)=0$；④ w 是语句 s 中的情感表达词语；⑤ $pol(w)$ 是 w 在语句 s 中的极性；⑥ $ntusd(w, pol(w))$ 是 w 基于 $pol(w)$ 的 NTUSD 分值。

由此，第 i 个用户的发文情感倾向 $ALen_i$ 为其在数据抓取期间，所有 n 个发文的情感平均值，如公式（7）所示。将 $ALen_i$ 标准化 [0,1] 后，作为第 i 个用户的发文情感值。

$$ASen_{i}=\frac{1}{n}\sum_{j=1}^{n}senti(p_{ij}) \qquad 公式 (7)$$

（2）用户发文的文本长度。第 i 个用户发文的文本长度 $ALen_i$ 为其在数据抓取期间，所有 n 个发文的文本长度平均值，如公式（8）所示：

$$ALen_{i}=\frac{1}{n}\sum_{j=1}^{n}lenth(p_{ij}) \qquad 公式 (8)$$

其中，$Length(p_{ij})$ 为第 i 个用户第 j 条发文 p_{ij} 的文本长度。将 $ALen_i$ 标准化 [0,1] 后，作为第 i 个用户发文的文本长度值。

（3）用户的社区等级。百度贴吧的等级制度是从 2011 年 9 月 20 日开始实行的一项制度，通过累计在线时长、参与社区讨论和每日签到等，得到用户的社区等级，在贴吧中，用户的社区等级分为 1-9，反映了用户的社区参与程度。第 i 个用户的社区等级分类值，如公式 (9) 所示：

$$AGrd_i=0 或 1 \qquad 公式（9）$$

其中：①当用户社区等级在 1-5 区间，由于所需经验值很少，新注册的用户可以在数天之内达到，不能确切反映用户的资历，因此，将第 i 个用户的社区等级分类赋值为 0；②当用户社区等级在 6-9 区间，需要一段时间累计经验值，因此，将第 i 个用户的社区等级分类赋值为 1。

知识管理论坛，2019(3):163-172
DOI：10.13266/j.issn.2095-5472.2019.017

根据公式（7）- 公式（9），节点属性的描述性统计如表 5 所示：

表 5 节点属性的描述性统计（N=2301）

节点属性	均值	标准差
用户发文的情感倾向	0.031	0.136
用户发文的文本长度	0.163	0.190
用户的社区等级	4.429	1.369

3.2 用户时序交互模式的 NATERGM 拟合

3.2.1 用户发文情感倾向的影响

百度糖尿病贴吧中用户发文情感倾向的 NATERGM 拟合结果，见表 6。用户发文情感倾向的 12 个假设中，有 4 个不成立，这是由于发文情感积极的用户更可能抢先连接其他用户，由此更可能被其他用户优先连接、传递引用和循环引用。

表 6 用户发文情感倾向的 NATERGM 拟合结果

交互模式		参数值	标准差	假设检验
弧		4.228***	0.056	
互惠性	反馈	1.683***	0.092	H1a 成立
	回复	2.017***	0.091	H2a 成立
k-in-star	抢先	0.195***	0.004	H3a 成立
	拖延	0.009	0.034	H4a 不成立
k-out-star	优先	0.139***	0.005	H5a 成立
	滞后	0.042	0.044	H6a 不成立
传递性	桥接	0.137	0.090	H7a 不成立
	追随	0.473***	0.071	H8a 成立
	引用	0.334***	0.076	H9a 成立
循环性	反向桥接	0.268	0.071	H10a 不成立
	反向追随	0.339***	0.045	H11a 成立
	反向引用	0.285***	0.044	H12a 成立

注：*** p<0.001；** p<0.01；* p<0.05

3.2.2 用户发文文本长度的影响

百度糖尿病贴吧中用户发文文本长度的 NATERGM 拟合结果，见表 7。用户发文文本长度的 12 个假设中，有 4 个不成立，这是由于发文文本较长的用户更可能拖延连接、桥接和反向桥接其他用户，由此更可能被其他用户传递引用和循环引用。

表 7 用户发文文本长度的 NATERGM 拟合结果

交互模式		参数值	标准差	假设检验
弧		3.241***	0.050	
互惠性	反馈	1.412***	0.103	H1b 成立
	回复	1.863***	0.102	H2b 成立
k-in-star	抢先	0.024	0.001	H3b 不成立
	拖延	0.195***	0.01	H4b 成立
k-out-star	优先	0.017	0.001	H5b 不成立
	滞后	0.231***	0.001	H6b 成立
传递性	桥接	0.813***	0.103	H7b 成立
	追随	0.117	0.084	H8b 不成立
	引用	0.391***	0.080	H9b 成立
循环性	反向桥接	1.079***	0.073	H10b 成立
	反向追随	0.071	0.046	H11b 不成立
	反向引用	0.192***	0.046	H12b 成立

注：*** p<0.001；** p<0.01；* p<0.05

3.2.3 用户社区等级的影响

百度糖尿病贴吧中用户社区等级的 NATERGM 拟合结果，见表 8。用户社区等级的 12 个假设中，有 5 个不成立，这是由于社区等级较高的用户更可能抢先连接、桥接、传递追随其他用户，由此更可能反向桥接和循环追随其他用户。

3.3 NATERGM 拟合度检验

拟合度检验采用度分布参数 [19]，包括：标准入度、标准出度、偏态入度和偏态出度，这是因为在社交媒体网络中，度分布最能体现用户如何连接以及如何参与交互。首先，对数据抓取期间的数据，根据公式（4）以周为单位进行拟合检验，分别针对每种交互模式 a，得到基于 TERGM 和 NATERGM 的绝对差异向量

$(AD_a^1, AD_a^2, \ldots, AD_a^{66})$；接着，以 TERGM 为参照对象，进行威尔科克森符号秩检验，结果见表 9。NATERGM 拟合误差显著性低于 TERGM 拟合误差，具有更好的拟合度。

表 8 用户社区等级的 NATERGM 拟合结果

交互模式		参数值	标准差	假设检验
弧		3.390***	0.050	
互惠性	反馈	1.770***	0.103	H1c 成立
	回复	1.490***	0.102	H2c 成立
k-in-star	抢先	0.110***	0.001	H3c 成立
	拖延	0.015	0.01	H4c 不成立
k-out-star	优先	0.075	0.001	H5c 不成立
	滞后	0.051	0.001	H6c 不成立
传递性	桥接	0.562***	0.103	H7c 成立
	追随	0.318**	0.084	H8c 成立
	引用	-0.160	0.080	H9c 不成立
循环性	反向桥接	0.341***	0.071	H10c 成立
	反向追随	0.235***	0.046	H11c 成立
	反向引用	0.198	0.050	H12c 不成立

注：*** p<0.001；** p<0.01；* p<0.05

表 9 威尔科克森符号秩检验结果

统计指标	标准入度	标准出度	偏态入度	偏态出度
统计值	5.98***	5.67***	4.45***	4.89***

注：***p<0.001；**p<0.01；*p<0.05

4 研究结论及社区发展建议

笔者基于 NATERGM，构建在线健康社区中用户时序交互模式的分析模型，将百度糖尿病贴吧作为实证研究对象，验证了应用 NATERGM 分析用户时序交互模式的有效性，揭示了用户属性对用户时序交互模式形成的影响 [19]。

4.1 互惠性时序模式的形成

互惠性时序假设中 H1a、H1b、H1c、H2a、H2b、H2c 均成立，意味着发文情感积极的用户、发文文本较长的用户以及社区等级较高的用户倾向于互惠性时序回复与互惠性时序反馈。

由此，建议在线健康社区提供类似于即时通讯的机制，方便用户直接交流，并且为发帖用户提供推荐列表，邀请社区内发文情感积极的用户、发文文本较长的用户以及社区等级较高的用户参与话题的讨论，以促进用户之间的活跃交互。

4.2 k-star 时序模式的形成

k-in-star 时序抢先假设中 H3a 和 H3c 均成立，H3b 不成立；k-in-star 时序拖延假设中 H4b 成立，H4a 和 H4c 均不成立。这意味着当用户收到来自贴吧中众多用户的解答时，来自于发

知识管理论坛，2019(3):163-172
DOI：10.13266/j.issn.2095-5472.2019.017

文情感积极的用户和社区等级较高的用户解答更可能较早，即这两类用户更倾向于及时回复解答；发文文本较长的用户倾向于延迟但详尽地回复解答。

k-out-star 时序优先假设中 H5a 成立，H5b 和 H5c 均不成立；k-out-star 时序滞后假设中 H6b 成立，H6a 和 H6c 均不成立。这意味着当众多用户寻求解答时，发文情感积极的用户更可能被优先回答，但对社区等级较高的用户并无优先性；发文文本较长的用户更可能被滞后回答，这是由于解读长文本需要一定的时间；在线健康社区用户更加关注情感支持和信息支持。

情感支持有助于构建活跃的社区氛围，信息支持直接反映了用户的表述水平和投入程度，由此，建议在线健康社区根据发帖内容的丰富程度，给予用户一定的回复奖励，以提高用户回复质量。

4.3 传递性时序模式与循环性时序模式的形成

传递性时序桥接假设中 H7b 和 H7c 均成立，H7a 不成立；传递性时序追随假设中 H8a 和 H8c 均成立，H8b 不成立；传递性引用假设中 H9a 和 H9b 均成立，H9c 不成立；循环性时序桥接假设中 H10b 和 H10c 均成立，H10a 不成立；循环性时序追随假设中 H11a 和 H11c 均成立，H11b 不成立；循环性时序引用假设中 H12a 和 H12b 均成立，H12c 不成立。意味着情感积极的用户更可能通过积极追随其他用户，进而被其他用户关注引用其发帖，但还不足以成为桥接中介；发文文本较长的用户更可能担当桥接并被其他用户引用其发帖，这是由于发文文本较长的用户发文信息量大，因而在用户交互过程中提供了重要的信息支持；社区等级较高的用户更可能担当桥接并追随其他用户，这是由于社区等级较高的用户经验丰富而善于追寻有效信息源。

信息支持与经验分享在交互网络中很大程度上发挥了中介作用，由此，建议在线健康社区为用户提供推荐列表，能充分利用发帖文本较长用户的影响力以及社区等级较高用户的自身经验为其他用户提供交互契机。

5 研究局限与展望

本研究在节点属性构建中，选取了用户发文情感倾向、发文文本长度和社区等级作为用户节点属性，未来研究需要提高节点属性的完备性，例如增加用户发文专业性。在研究节点属性的影响方面，探讨了单一节点属性对网络动态形成的影响，未来研究需要在现有 NATERGM 时序模式的基础上，扩展研究多种节点属性相互影响与作用下的网络动态形成机理。在实证对象选取中，将百度糖尿病吧作为实证社区，未来需要在更多的在线健康社区中进行实证研究，探究不同社区中用户交互时序模式的差异。

参考文献：

[1] 吴江，周露莎. 医疗信息资源跨地区流动：在线医疗社区优化医疗资源配置作用的研究 [J]. 情报科学 ,2017，4（1）：58-65.

[2] 龙天悦. 在线医疗社区的持续使用行为及其对医患关系影响研究 [D]. 合肥：合肥工业大学 ,2017.

[3] ZHAO K, GREER G E, YEN J, et al. Leader identification in an online health community for cancer survivors: a social network-based classification approach[J]. Information systems and e-Business management, 2015, 13(4): 629-645.

[4] 吴江，周露莎. 在线医疗社区中知识共享网络及知识互动行为研究 [J]. 信息资源管理学报 ,2017，35（3）：144-151.

[5] 翟羽佳，张鑫，王芳. 在线健康社区中的用户参与行为——以“百度戒烟吧”为例 [J]. 图书情报工作 ,2017(7): 75-82.

[6] STEWART S A, ABIDI S S R. Applying social network analysis to understand the knowledge sharing behaviour of practitioners in a clinical online discussion forum [J]. Journal of medical Internet research, 2012, 14(6): e170-e175.

[7] 刘璇，汪林威，李嘉，等. 在线健康社区中用户回帖行为影响机理研究 [J]. 管理科学 ,2017,1(30):62-72.

[8] WU B, Jiang S, Chen H C. Effects of individual motivations on support networks for uses and gratifications in online health forums[J].Social behavior



and personality, 2016, 44(2): 299–312.

[9] 张星 , 夏火松 , 陈星 , 等 . 在线健康社区中信息可信性的影响因素研究 [J]. 图书情报工作 ,2015(22):88-96.

[10] JIANG S, CHEN H C. NATERGM: A model for examining the role of nodal attributes in dynamic social media networks[J].IEEE transactions on knowledge and data engineering,2016,28(3): 729-740.

[11] LUO P, CHEN K, WU C H, et al. Exploring the social influence of multichannel access in an online health community [J]. Journal of the Association for Information Science and Technology, 2018, 69(1): 98-109.

[12] LARSEN-FREEMAN D. Adjusting expectations: the study of complexity, accuracy, and fluency in second language acquisition[J]. Applied linguistics, 2009,30(4): 579-589.

[13] 张敏 , 马臻 , 张艳 . 在线健康社区中用户主观知识隐藏行为的形成路径 [J]. 情报理论与实践 ,2018(10): 111-117.

[14] 范晓妞 , 艾时钟 . 在线医疗社区参与双方行为对知识交换效果影响的实证研究 [J]. 情报杂志 ,2016(7):173-178.

[15] JIANG S, GAO Q, CHEN H C. The roles of sharing, transfer, and public funding in nanotechnology knowledge diffusion networks[J]. Journal of the Association for Information Science and Technology, 2014,66（1）: 1017–1029.

[16] GOODREAU S M, MORRIS M. Birds of a feather, or friend of a friend? using exponential random graph models to investigate adolescent social networks [J]. Demography, 2009，46（1）:103-125.

[17] 张敏 , 刘雪瑞 , 张艳 . 在线健康社区用户诊疗信息求助行为——外部因素、个体动机与形成路径 [J]. 现代情报 ,2018(11):18-24.

[18] ZIMBRA D, CHEN H C. A stakeholder approach to stock prediction using finance social media [J]. IEEE intelligent systems, 2011, 26(6): 88-92.

[19] HANNEKE S, FU W, XING E P. Discrete temporal models of social networks [J]. Electronic journal of statistics, 2010, 4（1）: 585-605.

[20] GOH J M, GAO G D, AGARWAL R. The creation of social value: can an online health community reduce rural-urban health disparities [J]. MIS quarterly, 2016, 40(1): 247-263.

作者贡献说明：
吴　冰：全文研究及撰写；
彭　彧：数据抓取与拟合。

Research on the Temporal Interactive Mode Based on User Attributes in Online Health Community

Wu Bing　Peng Yu
School of Economics and Management, Tongji University, Shanghai 200092

Abstract: [Purpose/significance] Online health community has been developing rapidly, but it is difficult to reveal the inherent mechanism between network node attribute and user dynamic interaction mode, with the lack of research on node attribute-based dynamic networks. **[Method/process]** This paper constructed the research model for users temporal interactive mode of online health community by applying NATERGM（Node Attribute-based Temporal Exponential Random Graph Model）, taking the online health community as research object, and combining with the feature of the user node properties that including users emotional tendency, the post text and the user community level. Then, data from Baidu Diabetes community from October 2016 to February 2018, which involves 2301 users, 6045 posts and 9490 replies, was used for empirical research. **[Results/conclusion]** Results indicate that the validity of NATERGM in analyzing temporal interactive mode, and node attributes have significant effects on reciprocity temporal mode, k-star temporal mode, transitivity temporal mode and cyclicity temporal mode. Consequently, suggestions for the development of online healthy community construction are put forward.

Keywords: online health community　node attribute　user interactive mode　NATERGM